BLG 454E - Learning From Data

HW1

Team Name: Feeding with My Hands

Members: **Mehmet Barış Yaman 150130136**
Can Yılmaz Altıniğne 150130132

Q1) A secret government agency has developed a scanner which determines whether a person is a terrorist. The scanner is fairly reliable; 95% of all scanned terrorists are identifed as terrorists, and 95% of all upstanding citizens are identified as such. An informant tells the agency that exactly one passenger of 100 aboard an aeroplane in which you are seated is a terrorist. The agency decide to scan each passenger and the shifty looking man sitting next to you is the first to test positive. What are the chances that this man is a terrorist?

We need to find P(Man is terrorist | Test is positive).

We will use Bayes:

If a man is a terrorist, the probability of test give the correct answer is 0.95, probability of wrong answer is 0.05

P(Man is terrorist | Test is positive) = (P(Test is positive | Man is terrorist) * P(Man is terrorist)) / P(Test is positive)

P(Test is positive) can be written as sum of all P(Test is positive | Man's situation) * P(Man's situation)

So denominator becomes

(P(Test is positive | Man is terrorist) * P(Man is terrorist)) + (P(Test is positive | Man isn't terrorist) * P(Man isn't terrorist))

P(Test is positive | Man is terrorist) = 0.95 and P(Man is terrorist) = 0.01, multiplication of two = 0.0095

P(Test is positive | Man isn't terrorist) = 0.05 and P(Man is not terrorist) = 0.99, multiplication of two = 0.0495

So the denominator is 0.0495 + 0.0095 = 0.059

P(Test is positive | Man is terrorist) = 0.95 and P(Man is terrorist) = 0.01, multiplication of two = 0.0095

Finally, P(Man is terrorist | Test is positive) is 0.0095 / 0.059 = 0.16

Q2) For a novel input x, a predictive model of the class c is given by p(c=1|x) = 0.7, p(c=2|x) = 0.2, p(c=3|x) = 0.1. The corresponding utility matrix U(ctrue,cpred) has elements:

$$\begin{bmatrix} 5 & 3 & 1 \\ 0 & 4 & -2 \\ -3 & 0 & 10 \end{bmatrix}$$

In terms of maximal expected utility, which is the best decision to take ?

p(c=1 | x) -> probability of selecting c = 1

p(c=2 | x) -> probability of selecting c = 2

p(c=3 | x) -> probability of selecting c = 3

Expected utility of predicting c = 1 -> p(c=1|x) * 5 + p(c=1|x) *0 + p(c=1|x) *(-3) = 1.4

Expected utility of predicting c = 2 -> p(c=2|x) * 3 + p(c=2|x) *4 + p(c=2|x) *0 = 1.4

Expected utility of predicting c = 3 -> p(c=3|x) * 1 + p(c=3|x) *(-2) + p(c=3|x) *10 = 0.9

Therefore, selecting first and second columns should be the best selections. In other words, c=1 and c=2 are the best selections.

Q3-1pt) Suppose the weights of randomly selected American female college students are normally distributed with unknown mean μ and standard deviation σ. A random sample of 10 American female college students yielded the following weights (in pounds):

115 122 130 127 149 160 152 138 149 180

Based on the definitions given above, identify the likelihood function and the maximum likelihood estimator of μ *and* σ^2, the mean weight and variance of all American female college students. Using the given sample, find a maximum likelihood estimate of μ as well.

Probability density function $= P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$

Likelihood function is $\prod_{i=1}^{n} P(x_i ; \mu, \sigma)$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

For maximum likelihood estimator of μ, we need to find derivative of likelihood function w.r.t μ.

We will take log of each side and find log-likelihood

$$\mathcal{L}(\mu, \sigma^2) = \sum_{i=1}^{N} (\log A + B(x-\mu)^2) \qquad A = \frac{1}{\sqrt{2\pi\sigma^2}} \qquad B = -\frac{1}{2\sigma^2}$$

$$= \sum_{i=1}^{N} \log A + B \sum_{i=1}^{N} (x_i - \mu)^2$$

$$\frac{d}{d\mu} \mathcal{L}(\mu, \sigma^2) = \frac{d}{d\mu} \sum_{i=1}^{N} \log A + \frac{d}{d\mu} B \sum_{i=1}^{N} (x_i - \mu)^2 = 0$$

0

$$B\left[\frac{d}{d\mu}\sum_{i=1}^{N}(x_i-\mu)^2\right]=0$$

$$B\left[-2\sum_{i=1}^{N}(x_i-\mu)\right]=0$$

$$\sum_{i=1}^{N}x_i-\sum_{i=1}^{N}\mu=0 \longrightarrow \sum_{i=1}^{N}x_i=N\mu$$

$$\Rightarrow \hat{\mu}=\frac{1}{N}\sum_{i=1}^{N}x_i$$

For maximum likelihood estimator of σ^2, take derivative of L w.r.t σ^2 and set it to 0.

$$L(\mu,\sigma^2)=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n\cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2}$$

- Take log of each side

$$\mathcal{L}(\mu,\sigma^2)=\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n+\ln e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2}$$

log likelihood

$$\mathcal{L}(\mu,\sigma^2)=\ln(2\pi\sigma^2)^{-\frac{n}{2}}-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$=-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

- Take the derivative w.r.t σ^2, and set it to 0.

$$\frac{\partial\mathcal{L}}{\partial\sigma^2}=-\frac{n}{2}\left(\frac{2\pi}{2\pi\sigma^2}\right)+\frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2=0$$

$$-\frac{n}{2\sigma^2}+\frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2=0 \Longrightarrow -\frac{n}{2(\sigma^2)^2}\left(\sigma^2-\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2\right)=0$$

multiply by $2(\sigma^2)^2$

$$=-n\sigma^2+\sum_{i=1}^{n}(x_i-\mu)^2=0$$

$$\sigma^2=\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$$

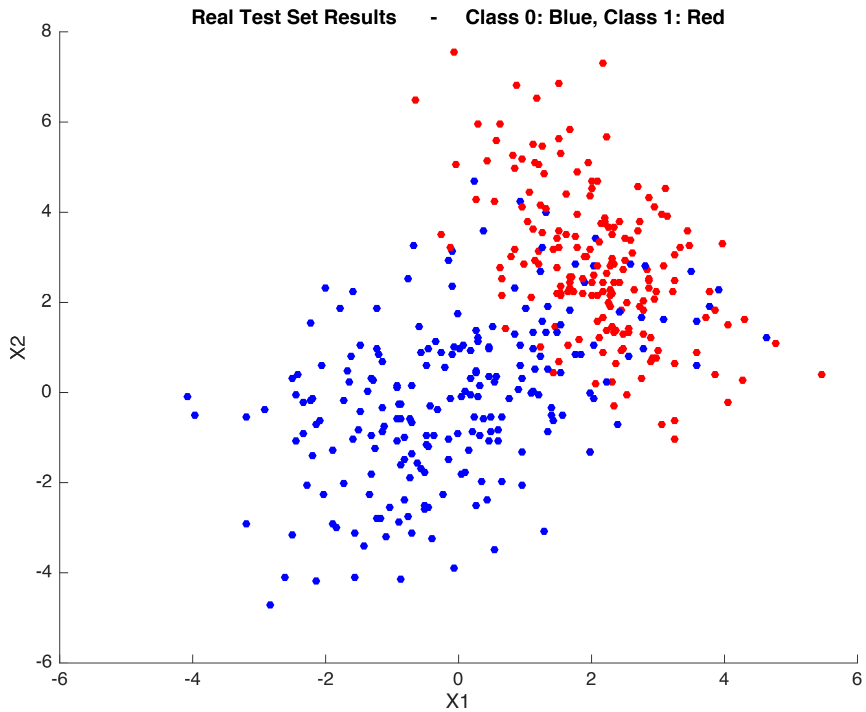
Maximum likelihood estimate of μ is 1/N * ∑ Xi as it is found.

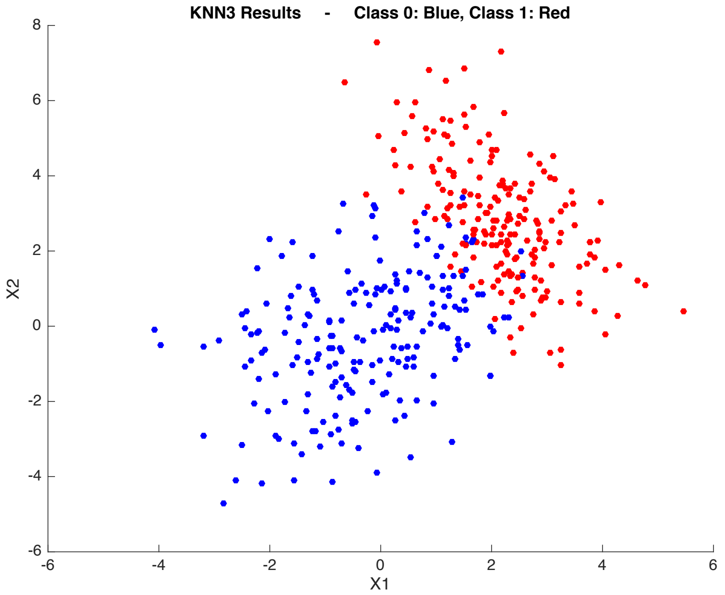
1/N = 1 / 10, ∑ Xi = 115+122+130+127+149+160+152+138+149+180 = 1422

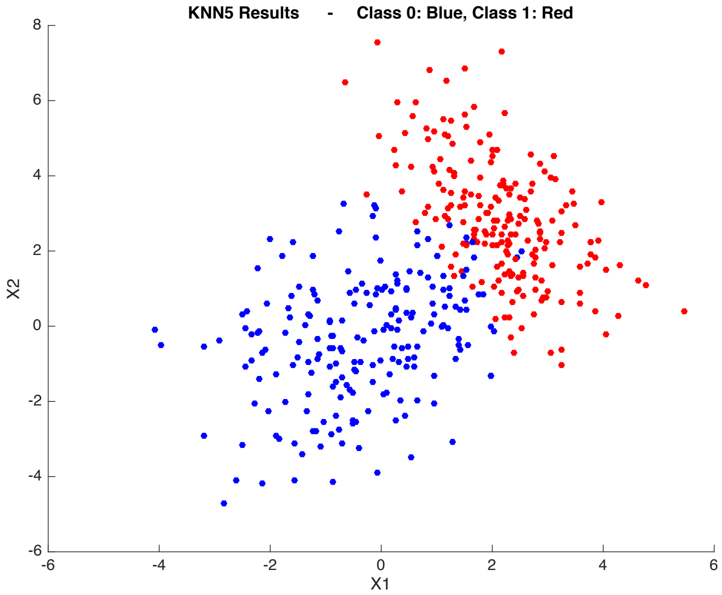
1/N * ∑ Xi = 142,2

Q4) Examine the dataset. The number of features and the number of classes. Classify given dataset using KNN classifier. Use at least two KNN (i.e KNN3 and KNN5) and compare the differences. Write comments below codes you implemented.

We randomly shuffled the dataset and take 80% of it for training set and 20% for test set. Then we found distance of every data in test set from every data in training set. Then we sorted the matrix which keeps distances for every data in test set from every data in training set. Then we took first K data from sorted distance matrix and check for 0/1 situation. If zeros are majority, then the predicted value for that test data becomes 0 else it becomes 1. We used KNN3 and KNN5 and got these results.







KNN3

Correct Predictions: 355

Wrong Predictions: 45

Accuracy: 88.75%

KNN5

Correct Predictions: 359

Wrong Predictions: 41

Accuracy: 89.75%

Generally we had accuracy around %87 and %91. Most of the time KNN5 has 1% - 2% more accuracy than KNN3 has.